

万卷方法

Microeconometrics Using Stata

用Stata学微观计量经济学

[美] A.科林·卡梅伦 [美] 普拉温·K.特里维迪 著
肖光恩 杨洋 王保双 等 译
肖光恩 审校



重庆大学出版社

<http://www.cqup.com.cn>

万卷方法

Microeconometrics Using Stata

用Stata学微观计量经济学

[美] A. 科林·卡梅伦 [美] 普拉温·K. 特里维迪 著
肖光恩 杨洋 王保双 等 译
肖光恩 审校



重庆大学出版社

Copyright © 2009, 2010 by StataCorp LP
All rights reserved. First edition 2009
Revised edition 2010

Published by Stata Press, 4905 Lakeway Drive, College Station, Texas 77845

No part of this book may be reproduced, stored in a retrieval system, or transcribed, in any form or by any means—electronic, mechanical, photocopy, recording, or otherwise—without the prior written permission of StataCorp LP.

Stata is a registered trademark of StataCorp LP. L^AT_EX 2 ϵ is a trademark of the American Mathematical Society.

《用 Stata 学微观计量经济学》原书英文版由 Stata 公司出版。原书版权属 Stata 公司。本书简体中文版权由 Stata 公司授予重庆大学出版社, 未经出版者书面许可, 不得以任何形式复制。

版贸核渝字(2013)第 206 号

图书在版编目(CIP)数据

用 Stata 学微观计量经济学/(美)卡梅伦
(Cameron, A. C.), (美)特里维迪(Trivedi, P. K.)著;
肖光恩, 杨洋, 王保双等译. —重庆: 重庆大学出版社,
2015. 4
(万卷方法)
ISBN 978-7-5624-8536-0
I. ①用… II. ①卡…②特…③肖…④杨…⑤王…
III. ①微观经济学—计量经济学—应用软件 IV.
①F016-39②F224. 0-39

中国版本图书馆 CIP 数据核字(2014)第 197484 号

用 Stata 学微观计量经济学

[美]A. 科林·卡梅伦 普拉温·K. 特里维迪 著
肖光恩 杨洋 王保双 等译
肖光恩 审校
策划编辑: 雷少波 林佳木 邹荣
责任编辑: 林佳木 邹荣 版式设计: 林佳木
责任校对: 关德强 责任印制: 赵晟

*

重庆大学出版社出版发行
出版人: 邓晓益
社址: 重庆市沙坪坝区大学城西路 21 号
邮编: 401331
电话: (023) 88617190 88617185(中小学)
传真: (023) 88617186 88617166
网址: <http://www.cqup.com.cn>
邮箱: fxk@cqup.com.cn (营销中心)

全国新华书店经销
自贡兴华印务有限公司印刷

*

开本: 787×1092 1/16 印张: 38.5 字数: 843 千
2015 年 4 月第 1 版 2015 年 4 月第 1 次印刷
印数: 1—4 000
ISBN 978-7-5624-8536-0 定价: 98.00 元

本书如有印刷、装订等质量问题, 本社负责调换
版权所有, 请勿擅自翻印和用本书
制作各类出版物及配套用书, 违者必究

Stata 专业技术人员对本书的评论

(<http://www.stata.com/bookstore/microeconometrics—stata/index.html>)

由美国学者科林·卡梅伦和普拉温·K. 特里维迪共同撰写的《用 Stata 学微观计量经济学》一书,是一本优秀的介绍微观计量经济学的专著,它同时也介绍了如何使用 Stata 来进行微观计量经济学研究。本书包括了微观计量经济学教材中省略的许多主题,同时也省略了对 Stata 基本使用知识的介绍。两位学者对 Stata 现有的微观计量经济学的方法进行了全面的和最新的总结。

本书的修订版更新了微观计量经济专家对 stata 11 版提出的新特征,本书不再使用 `mfx` 命令和用户编写的 `margeff` 命令,而是使用新的 `margins` 命令,它强调了在均值处的边际效应和平均的边际效应。让可以识别指示变量和交互效应的因子变量取代了 `xi` 命令,同时用几个实证例子说明了用于广义矩估计和非线性工具变量估计的新命令 `gmm`。最后,有关最大似然估计章节对 stata 11 中的命令 `ml` 的更新功能进行了介绍。

卡梅伦和特里维迪在本书中首先介绍了模拟的方法,然后用模拟的方法去演示了本书随后介绍的估计量和统计检验的特征。尽管模拟方法是微观计量经济学专家的重要工具,但它并不包含在标准的计量经济学教材中。通过对模拟方法的介绍,作者使用学生和研究者可以在未来的研究中使用这种武器。卡梅伦和特里维迪用详细的实例说明了每一个主题,而且在适当的时候他们也引用了 2005 年出版的《微观计量经济学:方法与应用》这本教材的内容。

而且两位作者还介绍了为执行那些没有 Stata 特定命令的计量分析方法而进行 stata 编程的特征。尽管本书不是定位于 stata 编程,但它也介绍了如何解决许多编程问题的方法。这些技术是实证计量经济学的基础,因为有许多新的特定的方法总会超出软件工具包现有的功能。

卡梅伦和特里维迪对主题的选择完全反映了当代微观计量经济学实证分析的操作,在向读者介绍了 Stata 之后,作者就介绍了线性回归、模拟和广义最小二乘法;有横截面技术的部分则介绍了对线性模型工具变量和分位数回归模型的最新处理方法。

本书后面的部分包括了对线性面板数据模型参数的估计方法,作者的选择是独具匠心的;在介绍了标准的随机效应和固定效应之后,然后介绍了混合线性模型,它在计量经济学之外的很多领域被广泛地应用。

卡梅伦和特里维迪不仅介绍了非线性回归模型,还说明了如何使用 Stata 对新的非线性估计量进行编程。此外,还详细地介绍了非线性的方法,这是很多传统计量经济学教材

省略的内容,这一部分向学生和研究者说明了如何更简单地执行这些新的非线性估计的方法。

然后作者使用经验的近似方法来说明统计推断,同时使用自抽样的近似方法来说明检验统计量的分布。这一部分强调了 Stata 在获取自抽样近似法方面的强大优势,同时也介绍了统计推断的基本知识。

最后,卡梅伦和特里维迪在扩展部分中介绍了不同的非线性模型,他们首先介绍了二值因变量模型,然后介绍了多项选择模型、tobit 和选择模型、计数模型以及非线性面板数据模型。最后用两个有关编程的附录结束了本书的写作。

本书对不同计量经济学主题的独特组合,对不同计量实证分析方法的直觉介绍,以及对 Stata 实例的深入解说,对使用微观计量经济学实证分析方法的人来说,它是最有价值的手边工具书。

推荐序:微观计量经济学实证分析的必备工具

当前计量经济学已经成为举足轻重的基础性学科,并成为理论研究和经验分析的重要工具。自美国两位经济学家詹姆士·J.赫克曼(James J. Heckman)和丹尼尔·L.麦克法登(Daniel L. McFadden)获得诺贝尔经济学奖以来,微观计量经济学与宏观计量经济学的区分就不断得到强化,特别是微观计量经济学理论得到了社会各界充分的重视,并取得了长足的发展。与此同时,随着美国与欧洲一些国家微观调查数据的不断涌现,特别是对大量个人、家庭或企业微观信息的强烈关注和追踪调查,微观计量分析已经成为现实生活与工作的重要内容。数据的实证分析者不仅需要微观计量经济学理论的强力支持,更需要有对微观大数据的处理能力和分析能力。因此,在加强学习计量经济学理论及其他社会经济学理论的同时,掌握对微观大数据的处理、加工、计算与分析技术,就成为当代社会各界的理论研究者与实证分析者的现实需要。

美国学者科林·卡梅伦和普拉温·K.特里维迪是当前世界公认的微观计量经济学专家,他们对微观大数据特别是对计数数据有专门的研究,他们编写的《微观计量经济学:方法与应用》一书更是首部以“微观计量经济学”命名计量经济学教材,这本教材的出版对计量经济学发展产生了重要影响。两位教授编写的《用 Stata 学微观计量经济学》一书,是近期得到迅猛发展的计量分析软件 Stata 所在公司量身打造的一本微观计量实证分析的专著。

这本书从实证分析者的视角,运用 Stata 软件分析的实例,深入浅出地讲解了当前微观计量经济学实证分析的许多重要方法。这些方法主要有模拟分析、工具变量回归、分位数回归、面板数据分析、非线性回归、非线性优化、自抽样、二值结果变量分析、多项选择模型、样本选择分析、计数分析、非线性面板分析等,本书介绍的这些微观实证分析方法基本涵盖了当前微观计量经济学领域的最新发展。与此同时,这本书还比较详细地介绍了在微观计量实证分析中所使用的 Stata 更新版本中的一些新命令的功能,它使计量经济学实证分析功能得到了极大的提升;更为重要的是本书还对 Stata 分析中所使用的矩阵编程语言 mata 进行了初步介绍,这对广大的 Stata 编程用户来说是一个很大的福利。因此,本书的翻译出版,不仅有利于推动 Stata 在中国的普及与推广,而且有利于广大计量经济分析爱好者提高实证分析能力。

《用 Stata 学微观计量经济学》一书是微观计量实证分析的必备工具,它有助于微观计量实证分析迈入大数据分析的时代。

连玉君

2014年6月于中山大学岭南学院

译者序

微观计量经济学是概率统计理论、计量经济学理论、微观数据处理和计量实证分析软件深度融合发展的结果。微观计量经济学理论的快速发展与广泛应用,不仅改变了传统计量经济学实证分析过度关注宏观经济加总数据和时间序列数据的分析,而且使时间序列数据与横截面数据嵌套构成的面板数据分析进入了新阶段。它把计量经济实证分析的重点聚焦在微观经济主体的同时,更强调微观主体个体特征对数据结构和计量方法的影响。在数据结构的构建上,大量微观数据的出现,使得微观计量分析从过度关注平衡面板数据分析逐渐转向非平衡面板数据和缺失数据的分析,而且在面板数据分析中从过度关注大 N 小 T 的面板数据逐渐转向小 N 大 T 的面板数据,更为重要的是从以前过度关注单层数据分析转向了多层数据分析,使分析的数据结构从单一水平结构转向了多层水平结构。在计量分析方法上,随着 Melitz(2003)异质性企业理论的广泛应用,很多计量经济实证分析者已经放松了过去有关观测单位是同质性且不相关的假设,并在实证计量分析中更关注观测单位的相关性与异质性特征对计量实证分析带来的影响,如对数据聚类性和分层性的控制以及对数据时间相关(或滞后)、空间相关(或滞后)或数据的时间-空间相关的控制等;与此同时,微观计量实证分析更重视微观观测单位个体特征与宏观数据的交互影响,使分层数据的计量实证分析方法不断丰富。因此,微观计量实证分析使理论与现实客观世界的联系更加紧密。

《用 Stata 学微观计量经济学》是一本计量经济学理论分析与实证分析深入结合的“手册式”工具书,本书的作者是当今世界研究微观计量经济学的重要权威,他们不仅对微观计量经济学最新的发展做出了巨大贡献,而且对 Stata 软件的编程分析具有重要的建树。本书是计量经济学理论研究与实证分析结合得十分完美的著作,是实证微观计量分析的重要工具书。本书对微观计量经济学的最新发展进行了总结,它不仅涵盖了分析位数回归、线性工具变量回归、非线性回归、非线性优化方法、非线性面板模型、二值结果变量分析、多项选择模型、样本选择模型和计数模型的基本分析方法,而且还囊括了模拟分析、自抽样和各种计量经济学分析的检验方法。本书的另一个重要特征就是把非传统计量经济学的分析方法与计量分析编程结合起来,使得实证计量分析的爱好者能学习到大量的编程知识与技巧,并有可能开发出属于自己的计量分析程序。

出于对微观计量分析和 Stata 软件的爱,让我有足够的勇气与重庆大学出版社、Stata

出版公司和本书作者进行联系。2011年在重庆大学出版社向林佳木编辑进行了推荐,然后重庆大学出版社与Stata公司经过漫长的曲折谈判,终于在2013年签订了中文版翻译合同。本书的翻译既是我长期学习Stata软件的一个成果,也是我带领我指导的研究生学习计量经济学和Stata软件操作的集体成果。本书的初译分工如下:冯楠(第1章)、肖光恩和冯楠(第2章)、范思齐和肖光恩(第3章)、杨洋(第4章、第5章、第18章和附录)、汪艺(第6章)、肖光恩和冯楠(第7章)、王保双(第8章和第9章)、袁盼盼(第10章至第13章)和黄艳(第14章至第17章),李莎莎参加了序言和主题词索引初译的工作。全书的最终校译由肖光恩完成,杨洋负责书中部分译文和书中图表及程序运行结果的编制工作。

本书翻译用了近两年的时间,仅最终校译工作就经历了四个多月。本书翻译的完成,不仅要感谢研究生们的辛苦工作,更要特别感谢研究生杨洋陪我一起牺牲了4个多月来的所有周末时间。本书的翻译还得益于武汉大学哲学社会科学优势和特色学术领域建设计划项目“后危机时代世界经济格局变动对中国的机遇和挑战”和武汉大学“70后”创新团队,以及武汉大学经济与管理学院“计量经济学”平台课程等项目经费的支持。最后,要感谢重庆大学出版社林佳木编辑的认真工作,也要感谢我的家人对翻译工作的全力支持。

由于水平所限,错误与疏漏之处在所难免,敬请读者批评指正,以便及时修正。联系方式为:xiaoguangen@whu.edu.cn。

肖光恩

武汉大学经济与管理学院

2014年8月3日

修订版序

2008年12月出版的《用 Stata 学习微观计量经济学》主要是针对 Stata 10.1 而写的。这本书包含了 Stata 10.0 的升级版 Stata 10.1 的新增内容,最重要的就是新的随机数生成器。

在修订版中,除了 Stata 10 的内容之外,我们介绍了首次出现在 Stata 11 中的内容。除了极少数的内容之外,我们用一种能重现在第一版中所出现的结果的方法来介绍这些新增的内容。

第一,我们介绍了因子变量新的构造方法。这些方法对设定具有一系列形成于分类变量的指示变量的模型和设定具有交互项的模型提供了一种简单的方法。因子变量取代了 xi 前缀命令。具体内容见 1.3.4 节和 2.4.7 节末尾。

第二,我们介绍了用于在回归模型中进行预测和对边际效应进行计算的新的 margins 命令。用带有包括 dydx() 选项的 margins 命令取代了 Stata 中的 mfx 命令以及用户编写的 margeff 命令。此外,在结合因子变量使用 margins 命令时,它可以简化具有交互项的模型中边际效应的计算。见 10.5 节和 10.6 节,特别是 10.5.7 节和 10.6.5 节。在整个修订版中,特别是在 14—17 章中,我们用 margins 命令取代了 mfx 和 margeff 命令。

在第一版中,我们通常计算在均值处的边际效应(MEM),而不是平均边际效应(AME),因为 mfx 命令不能计算 AME。新的 margins 命令既可以计算 MEM 也可以计算 AME。在修订版中,我们努力重现第一版中已经给出的分析结果。因为这个原因,我们还将继续更频繁地计算 MEM,尽管在实际应用中,AME 更受欢迎。

第三,我们介绍了用于广义矩方法和非线性工具变量估计方法中的新 gmm 命令。见 10.3.8 节和 17.5.2 节。

第四,我们介绍了在使用 d1 和 d2 方法时,对已有的 ml 命令所做的一些细小改变。这些改变的出现是因为现在的 ml 命令是新的 Mata moptimize() 函数的前端命令。同时,我们也介绍了新的 lf0、lf1 和 lf2 方法,见 11.6 节。Mata optimize() v 计算器已经被更名为 gf 计算器,见 11.7 节。

我们感谢 Stata 公司的全体员工在修订版的准备工作中所提供的帮助,尤其是 Patricia Branton、David Drukker、Lisa Gilmore 以及 Deirdre Patterson。

加利福尼亚州 戴维斯 A. 科林·卡梅伦
印第安纳州布卢明顿 普拉温·K. 特里维迪

2010年1月

第一版序

本书解释了如何使用计量经济学计算机软件 Stata 来执行对横截面和面板数据进行回归分析。书名中使用的微观计量经济学这个术语,是因为其应用于与经济学相关的数据,并且因为其内容包含了诸如工具变量回归的方法,与在应用统计学的其他领域中的应用相比,此方法在经济学领域中运用得更多。然而,本书中所讨论的许多主题、模型和方法也与其他社会科学是相关的。

本书的主要读者是研究生和研究人员。对他们来说,可以把这本书作为我们自己的著作《微观计量经济学:方法与应用》(Cameron and Trivedi 2005),以及其他研究生水平的教材如 Greene(2008) 和 Wooldridge(2002)等的补充读物。与这些书相比,本书并不介绍计量经济学理论,相反,本书强调使用 Stata 来执行某些实际操作。我们介绍了更多前沿的主题,包括分位数回归、弱工具变量、非线性优化法、自抽样法、非线性面板数据分析法以及 Stata 的矩阵编程语言 Mata。

与此同时,本书还介绍了诸如广义最小二乘回归、工具变量估计、logit 模型和 probit 模型的内容,它可以作为合适的本科生水平的计量经济学教材的补充读物,因此它非常适合在本科生的计量经济学课堂中使用。下表建议本书中的这些章节可以作为计量经济学导论的内容,需要注意的是这些章节中出现的某些公式使用了矩阵代数。

Stata 基础知识	第 1.1—1.4 节
数据管理	第 2.1—2.4, 2.6 节
OLS	第 3.1—3.6 节
模拟	第 4.6—4.7 节
GLS(异方差)	第 5.3 节
工具变量	第 6.2—6.3 节
线性面板数据	第 8 章
logit 和 probit 模型	第 14.1—14.4 节
tobit 模型	第 16.1—16.3 节

虽然我们提供了有关 Stata 的很多细节,但是这种处理方法并不完整,特别是,我们介绍了各种不同的 Stata 命令,但没有对这些命令进行详细的列表和描述,因为它们 Stata 手册和网络在线帮助中已经很好地进行了说明。通常,我们只给出一个提示、一个简短的

讨论,或者只给出一个例子。

我们尽可能提供一个模板程序,其可以被改编而适用于其他问题。请记住,为了减少本书中的输出结果,我们的例子中使用的解释变量比真正的严谨的研究所必需的解释变量要少得多。我们的程序经常压缩在实际研究中很重要的中间输出结果,因为广泛使用了 `quietly` 命令和 `nolog`、`nodots`、`noheader` 选项。并且我们还尽量避免使用在探索性数据分析中普遍使用的图表。

我们已经使用了 Stata 10 以及 Stata 的更新内容。^① 关于如何获得本书中使用的数据集和 do 文件的介绍,可以在 Stata 出版社的网页 <http://www.stata-press.com/data/mus.html> 中获得。对本书进行的所有修正的文件可以在 <http://www.stata-press.com/booksfmus.html> 中获得。

在这个项目中,除了学习 Stata 之外,我们还学了很多计量经济学知识。事实上,我们坚信,学习计量经济学的一个有效方式就是干中学,即打开一个 Stata 数据集,查看使用不同方法以及这些方法的变化所得到的结果,比如使用稳健标准误,而不是使用默认标准误。这种方法对学习计量经济学的各种能力都是有益的。实际上,能让你熟悉 Stata 的主要特征的一个有效方式可能是用你自己的数据集执行相关章节中的命令。

我们感谢在本书的准备过程中许多提供过帮助的人。这个项目使我们出版了 2005 年的教材,并且我们感谢 Scott Parris 对那本书的专业处理。Juan Du、Qian Li 和 Abhijit Ramalingam 仔细阅读了本书的许多章节。与 John Daniels、Oscar Jorda、Guido Kuersteiner 和 Doug Miller 进行的讨论也是特别有帮助的。感谢 Deirdre Patterson 优秀的编辑和 Lisa Gilmore 对本书的 L^AT_EX 排版和本书的出版,尤其要感谢 David Drukker 在这个项目的各个阶段所给予的大量投入和鼓励,包括对最终稿的深入阅读和评论,这使得本书在计量经济学和 Stata 的相关内容上有了许多改进。最后,感谢我们各自的家庭所做出的不可避免的牺牲,以至于我们能开展并完成了这个多年的项目。

加利福尼亚州戴维斯 A. 科林·卡梅伦
印第安纳州布卢明顿 普拉温·K. 特里维迪
2008 年 10 月

① 为了查看你是否有最新的更新,输入 `update query` 命令。对于使用 Stata 早期版本的用户,一些关键的变化如下:Stata 9 引入了矩阵编程语言 Mata。Stata 10 的语法使用了 `vce(robust)` 选项而不是 `robust` 选项来获得稳健标准误。在 2008 年中,在 Stata 10 的更新中引入了新的随机数函数,例如 `runiform()` 和 `rnormal()`。

目 录

1 Stata 基础知识	1
1.1 Stata 的交互使用	1
1.2 Stata 文件	2
1.2.1 Stata 手册	2
1.2.2 其他 Stata 资源	2
1.2.3 help 命令	3
1.2.4 search、findit 和 hsearch 命令	4
1.3 命令的语法和算子	4
1.3.1 基本的命令语法	4
1.3.2 例子: summarize 命令	5
1.3.3 例子: regress 命令	6
1.3.4 因子变量	7
1.3.5 缩写词、大小写的敏感性和通配符	9
1.3.6 算术、关系和逻辑算子	10
1.3.7 报错的信息	10
1.4 do 文件和 log 文件	11
1.4.1 编写一个 do 文件	11
1.4.2 运行 do 文件	11
1.4.3 log 文件	12
1.4.4 一个三阶段的过程	13
1.4.5 注释和长行的处理	13
1.4.6 执行 Stata 的不同方法	14
1.5 标量和矩阵	14
1.5.1 标量	14
1.5.2 矩阵	15
1.6 使用来自于 Stata 命令后所获得的结果	15
1.6.1 使用从 r 类命令 summarize 所得的结果	15
1.6.2 使用从 e 类命令 regress 所得的结果	16
1.7 全局宏和局部宏	18
1.7.1 全局宏	18
1.7.2 局部宏	18
1.7.3 使用标量还是使用宏?	19
1.8 循环命令	20

1.8.1	foreach 循环	21
1.8.2	forvalues 循环	21
1.8.3	while 循环	22
1.8.4	continue 命令	22
1.9	一些有用的命令	22
1.10	do 文件的模板	22
1.11	用户编写的命令	23
1.12	Stata 资源	24
1.13	习题	24
2	数据管理和绘图	25
2.1	导论	25
2.2	数据的类型	25
2.2.1	文本或 ASCII 数据	25
2.2.2	内部数字数据	26
2.2.3	字符串数据	27
2.2.4	数字型数据的显示格式	27
2.3	输入数据	27
2.3.1	基本原则	27
2.3.2	输入 Stata 格式的数据	28
2.3.3	用键盘输入数据	29
2.3.4	输入非文本数据	29
2.3.5	从电子表格中输入文本数据	30
2.3.6	输入自由格式的文本数据	31
2.3.7	输入固定格式的文本数据	31
2.3.8	字典数据文件	32
2.3.9	常见的误区	32
2.4	数据管理	32
2.4.1	PSID 的例子	32
2.4.2	对变量命名和贴标签	34
2.4.3	查看数据	35
2.4.4	使用原始文档	36
2.4.5	缺失值	36
2.4.6	估算缺失的数据	38
2.4.7	转换数据 (generate、replace、egen 和 recode 命令)	38
2.4.8	保存数据	42
2.4.9	选择样本	43
2.5	操作数据集	43
2.5.1	观测值和变量的排序	44
2.5.2	保存和恢复数据集	44
2.5.3	数据集的宽格式和长格式	44
2.5.4	合并数据集	44

2.5.5	追加数据集	46
2.6	数据的图形显示	47
2.6.1	Stata 的图形命令	47
2.6.2	箱线图	49
2.6.3	直方图	50
2.6.4	核密度图	51
2.6.5	二维散点图和拟合线	52
2.6.6	局部加权散点平滑估计、核密度、局部线性和最近邻回归	53
2.6.7	多重散点图	55
2.7	Stata 资源	56
2.8	习题	57
3	线性回归的基本知识	59
3.1	导论	59
3.2	数据与数据的概述性统计	59
3.2.1	数据的描述	59
3.2.2	变量的描述	60
3.2.3	概述性统计	60
3.2.4	更详细的概述性统计	61
3.2.5	数据的统计表	62
3.2.6	统计检验	64
3.2.7	数据的绘图	64
3.3	水平值和对数值的回归	65
3.3.1	基本的回归理论	65
3.3.2	OLS 回归和矩阵的代数运算	66
3.3.3	OLS 估计量的性质	66
3.3.4	异方差-稳健标准误	67
3.3.5	聚类-稳健标准误	67
3.3.6	对数值的回归	68
3.4	基本的回归分析	69
3.4.1	相关性	69
3.4.2	regress 命令	69
3.4.3	假设检验	71
3.4.4	多个回归输出结果的合并表	71
3.4.5	更好的回归输出表	72
3.4.6	用于生成分类变量和交互项的因子变量	74
3.5	模型设定的分析	75
3.5.1	模型设定的检验和模型的诊断	75
3.5.2	残差的诊断图	76
3.5.3	有影响力的观测值	77
3.5.4	模型设定的检验	78
3.5.5	具有多种幂形式的模型检验	81

3.6	预测	83
3.6.1	样本内预测	83
3.6.2	ME 和弹性	84
3.6.3	对数形式的预测:再转换的问题	86
3.6.4	预测的练习	87
3.7	抽样权重	88
3.7.1	权重	88
3.7.2	加权的均值	89
3.7.3	加权的回归	90
3.7.4	加权的预测和 ME	91
3.8	运用 Mata 进行 OLS 回归	92
3.9	Stata 资源	93
3.10	习题	93
4	模拟	95
4.1	导论	95
4.2	伪随机数生成器:导论	95
4.2.1	均匀分布随机数的生成	96
4.2.2	从正态分布中抽样	97
4.2.3	从 t 分布、卡方分布、 F 分布、 γ 分布和 β 分布中进行抽样	98
4.2.4	从二项分布、泊松分布和负二项分布中进行抽样	99
4.3	样本均值的分布	101
4.3.1	Stata 程序	102
4.3.2	simulate 命令	103
4.3.3	中心极限定理的模拟	103
4.3.4	postfile 命令	104
4.3.5	模拟中心极限定理的另一种方法	104
4.4	伪随机数生成器:更详细的介绍	105
4.4.1	逆概率转换法	105
4.4.2	直接转换法	107
4.4.3	其他方法	107
4.4.4	从截尾正态分布中进行抽样	107
4.4.5	从多元正态分布中进行抽样	108
4.4.6	使用马尔科夫链蒙特卡洛方法进行抽样	109
4.5	计算积分	111
4.5.1	求积法	111
4.5.2	蒙特卡洛积分法	111
4.5.3	使用不同 S 值的蒙特卡洛积分	112
4.6	回归的模拟:导论	113
4.6.1	模拟的例子:具有卡方分布误差的 OLS	113
4.6.2	解释模拟结果	115
4.6.3	模拟的差异	117

4.6.4	估计量的非一致性	118
4.6.5	具有内生解释变量的模拟	119
4.7	Stata 资源	120
4.8	习题	120
5	GLS 回归	123
5.1	导论	123
5.2	GLS 回归和 FGLS 回归	123
5.2.1	用于异方差误差项的 GLS	123
5.2.2	GLS 与 FGLS	124
5.2.3	加权最小二乘法和稳健标准误	124
5.2.4	重要的例子	125
5.3	对异方差数据进行建模	125
5.3.1	模拟的数据集	126
5.3.2	OLS 估计	126
5.3.3	检验异方差	127
5.3.4	FGLS 估计	129
5.3.5	WLS 估计	130
5.4	系统线性回归	130
5.4.1	SUR 模型	131
5.4.2	sureg 命令	132
5.4.3	在两类支出中的应用	132
5.4.4	稳健标准误	133
5.4.5	跨方程约束的检验	134
5.4.6	施加跨方程约束	135
5.5	调查数据:加权、聚类和分层	136
5.5.1	调查设计	136
5.5.2	调查均值的估计	138
5.5.3	调查数据的线性回归	139
5.6	Stata 资源	140
5.7	习题	140
6	线性工具变量回归	143
6.1	导论	143
6.2	IV 估计	143
6.2.1	基本的 IV 理论	143
6.2.2	模型的建立	144
6.2.3	IV 估计量: IV、2SLS 和 GMM	145
6.2.4	工具变量的有效性和相关性	146
6.2.5	稳健标准误的估计值	147
6.3	IV 的例子	147
6.3.1	ivregress 命令	148
6.3.2	具有一个内生性解释变量的医疗支出模型	148

6.3.3	可用的工具变量	149
6.3.4	恰好识别模型的 IV 估计	150
6.3.5	过度识别模型的 IV 估计	151
6.3.6	解释变量内生性的检验	152
6.3.7	过度识别约束的检验	154
6.3.8	二值内生性解释变量的 IV 估计	155
6.4	弱工具变量	157
6.4.1	IV 估计量的有限样本性质	157
6.4.2	弱工具变量	157
6.4.3	estat firststage 命令	159
6.4.4	恰好识别模型	159
6.4.5	过度识别模型	161
6.4.6	多个内生性解释变量	161
6.4.7	工具变量选择的敏感性分析	161
6.5	对弱工具变量的更好推断	163
6.5.1	条件检验和置信区间	163
6.5.2	LIML 估计量	164
6.5.3	刀切法 IV 估计量	165
6.5.4	2SLS、LIML、JIVE 和 GMM 的比较	165
6.6	3SLS 系统估计	166
6.7	Stata 的资源	168
6.8	习题	168
7	分位数回归	171
7.1	导论	171
7.2	分位数回归	171
7.2.1	条件分位数	172
7.2.2	QR 估计值和标准误的计算	172
7.2.3	qreg, bsqreg 和 sqreg 命令	173
7.3	医疗支出数据的分位数回归	174
7.3.1	数据的概述	174
7.3.2	QR 估计	174
7.3.3	条件分位数系数的解释	175
7.3.4	再转换	176
7.3.5	不同分位数上的估计值的比较	177
7.3.6	异方差的检验	178
7.3.7	假设检验	178
7.3.8	不同分位数回归中系数的图形显示	179
7.4	生成异方差数据的 QR	181
7.4.1	模拟产生的数据集	181
7.4.2	QR 的估计值	182
7.5	计数数据的 QR	184
7.5.1	分位数计数回归	184
7.5.2	qcount 命令	185